Center for Independent Experts (CIE) Reviewer's Independent Peer Review Report on the 62nd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC) Benchmark stock assessment for Black sea bass and Witch flounder

Prepared by

Neil Klaer

Prepared for

The Center for Independent Experts

Review Meeting
Woods Hole, Massachusetts
29 November – 2 December 2017

Contents

a. Provide numerical annual projections (3-5 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level) that fully incorporates observation, process and model uncertainty (see Appendix to the SAW TORs). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and definition of BRPs for black sea bass).
b. Comment on which projections seem most realistic. Consider major uncertainties in the assessment as well as the sensitivity of the projections to various assumptions.
c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

2.1.10. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

2.2 Review findings by term of reference for Witch flounder

2.2.1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

2.2.2. Present available federal, state, and other survey data, indices of relative or absolute abundance, recruitment, etc. Characterize the uncertainty and any bias in these sources of data and compare survey coverage to locations of fishery catches. Select the surveys and indices for use in the assessment.

2.2.3. Investigate effects of environmental factors and climate change on recruitment, growth and natural mortality of witch flounder. If quantifiable relationships are identified, consider incorporating these into the stock assessment.

2.2.4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3 if appropriate), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections. Compare F's and SSB's that were projected during the previous assessment to their realized values.

2.2.5. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

2.2.6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model (or possibly models, in accord with guidance in attached "Appendix to the SAW Assessment TORs") developed for this peer review. In both cases, evaluate whether the stock is rebuilt.
a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the updated BRP estimates.
b. Then use the newly proposed model (or possibly models, in accord with guidance in "Appendix to the SAW Assessment TORs") and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).

2.2.7. Develop approaches and apply them to conduct stock projections.

   a. Provide numerical annual projections (3 years) and the statistical distribution (e.g., probability density function) of the catch at FMSY or an FMSY proxy (i.e. the overfishing level, OFL) (see Appendix). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, magnitude and variability in recruitment).
   b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions. Identify reasonable projection parameters (recruitment, weight-at- age, retrospective adjustments, etc.) to use when setting specifications.
   c. Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC. The choice takes scientific uncertainty into account (see Appendix).

2.2.8. Evaluate the validity of the current stock definition, taking into account what is known about migration, and make a recommendation about whether there is a need to modify the current stock definition for future stock assessments.

2.2.9. Review, evaluate and report on the status of research recommendations from the last peer reviewed benchmark stock assessment. Identify new research recommendations.

Annex 1: Bibliography of materials provided for review
Annex 2: A copy of the CIE Statement of Work
Annex 3: List of participants

Executive Summary

The 62nd Stock Assessment Review Committee (SARC) on assessments of Black sea bass (*Centropristis striata*) and Witch flounder (*Glyptocephalus cynoglossus*) met in the Aquarium Conference Room at NOAA's Northeast Fisheries Science Center (NEFSC) in Woods Hole from Tuesday November 29 to Friday December 2, 2016. The meeting was chaired by Dr Patrick J. Sullivan from the New England Fisheries Management Council's Scientific and Statistical Committee and the review panel (the Panel) was composed of three scientists appointed by the Center for Independent Experts: Vivian Haist, Neil Klaer, and Anders Nielsen. The meeting generally followed the draft agenda and included presentations by representatives of the stock assessment teams: Dr Gary Shepherd for Black sea bass and Dr Susan Wigley for Witch flounder. The meeting was open to the public and was available also on live audio via the web. Discussion was mostly restricted to clarification during presentations with broader discussion at the conclusion of presentations. There was limited time available for additional work outside the meeting, but several homework analyses were requested by the Panel and results were provided during the meeting. The meeting atmosphere was congenial, the venue was excellent, and contributors to the meeting all did so in a professional and efficient manner, allowing the meeting work to be completed on-time and without contention.

Findings for Black sea bass

A Mid-Atlantic Fisheries Management Council Scientific and Statistical Committee peer review approved the SAW Working Group recommended spatial partitioning of the black sea bass population north of Cape Hatteras into North and South sub-units using the Hudson Canyon as the boundary. The Panel agrees that evidence, particularly from tagging and observed differences in recruitment patterns within sub-units provides a good basis for this choice of sub-units.

There were several fishery components comprising both the commercial (bottom trawl, handline, pot and other) and the recreational fisheries and each was well described. Methods used for the calculation of discards and associated errors seem appropriate. Further work is required to characterize catch/discard uncertainty – either by assignment of error to each (at least annually), or through construction of alternative matrices that provide scenarios that account for that error for stock assessment sensitivity testing.

Data filtering methods and index standardization procedures were not examined in detail for this review, but the information provided suggests that best-practice methods for these could be considered for review at a national level. Data collection associated with the fishery-independent trawl indices provide a good source of length and age data, and lengths are available from all surveys. There are very obvious differences in abundance trends in the north and south sub-units further justifying an area-based approach.

Further work is encouraged to determine environmental factors affecting abundance of different life stages – particularly in an ecosystem context for the NW Atlantic region. Broad patterns of change across species complexes require characterization that may lead to co-variates useful to stock assessments. Apparent northward movement of the center of population abundance is consistent with range movement associated with climate change as observed globally, so further work on range shift across species in the region and implications for stock assessment also seems warranted.

Application of the ASAP model to each sub-unit was largely successful (including an account of catch exchange between sub-units), except that strong and opposing retrospective patterns are evident within each sub-unit. To examine the entire stock, results from the assessed sub-units are added together, and when this is done, some retrospective pattern remains, although not sufficient to move the current stock status out of the not overfished and not overfishing quadrant after application of a rho adjustment. The Panel accepted the SAW Working Group's proposed base case and concurred that it provided a credible basis for providing management advice.

Although presented as a preliminary model for information only, an SS3 model was constructed that allowed partitioning of annual recruitment levels differently to each sub-area annually, and potentially allowed for mixing among areas. This model did not exhibit a substantial retrospective pattern that would require rho adjustment. A comparison of overall SSB and recruitment trends of the SS3 model and the combined ASAP results for the north and south sub-units showed overall agreement.

Biological reference points were calculated using results of the two-area stock assessments. The Panel agrees the BRPs calculated for black sea bass are appropriate. The Panel supports the conclusion of the SAW that the Black sea bass stock is not overfished and overfishing is not occurring.

Findings for Witch flounder

Methods for calculation of discards and associated errors seem appropriate. The magnitude of under-reported catch throughout the time series should be more fully assessed and documented. Using age-length keys on landings data from survey sources rather than those derived from landings age-length samples directly is problematic. If age samples are not available, it would be preferable to fit directly to the available length data by source.

The spring and autumn NEFSC surveys are regarded as providing the best available fishery independent indices for this species and they show broadly similar patterns of a decline from the early 1960s to record low levels in the late 1980s and early 1990s, an increase to early 2000s followed by a declining trend. Dealer Report LPUE, Vessel Trip Report, Observer Program and Study Fleet were evaluated as potential fishery-dependent indices. The Panel agrees with the evaluations made, and the provision of the LPUE index using 40% filtering as an appropriate best candidate. The Panel also agrees that quantification of bias in the LPUE index is a difficult technical problem and that quantitative measures of some important influencing factors may not be available.

Data filtering methods and index standardization procedures were not examined in detail for this review, but information provided suggests that best-practice methods for these could be considered for review at a national level.

The sweep study provides a q estimate that can be applied to the NEFSC survey to determine absolute abundance, and the Panel agrees that this is an appropriate piece of information to use for constructing an abundance series for this species.

The Panel rejected the SAW's ASAP base model because the major retrospective pattern exhibited by the stock reconstruction was deemed unacceptable, and none of the sensitivity runs presented were considered to be an acceptable alternative as they also had major retrospective patterns or reflected unacceptable assumptions. Additionally, some of the alternative models had quite different abundance estimates or trends than the base model indicating results were not robust to the uncertain assumptions of the analysis.

Retrospective analysis is a good method for identifying model misspecification (as caused by inconsistencies in data or model structure), and the apparent data quality issues or model misspecification in the witch flounder assessment needs to be resolved (e.g., applying a model or multiple models that remove the pattern while using an objective basis for model selection). Rho-based bias adjustment is an ad hoc procedure that may not correct such retrospective problems in the long term and as such it may not always provide appropriate management advice.

The Panel believes that the previously accepted VPA model is not an acceptable alternative to the rejected ASAP application because it exhibits a similar major retrospective pattern and the Panel recommends that the NEFMC SSC consider using the empirical approach discussed in the assessment document for use as the basis for developing management advice. The empirical analysis indicates stock biomass declined since 2002 although it appears to have stabilized in recent years.

The empirical approach, based on the NEFSCs spring and autumn surveys, calculates swept-area abundance incorporating results of the swept-area experiment. While absolute biomass estimates from this method will be highly uncertain (because of assumptions required in the conversion of survey catch rate to absolute estimates), the estimates of relative exploitation rates will be more robust to those uncertainties.

The alternative empirical approach was used to determined biomass and exploitation rates, but not biological reference points. As a consequence, stock status is currently unknown.

1 Introduction

1.1 Background

The 62nd Stock Assessment Review Committee (SARC) on assessments of Black sea bass (*Centropristis striata*) and Witch flounder (*Glyptocephalus cynoglossus*) met in the Aquarium Conference Room at NOAA's Northeast Fisheries Science Center (NEFSC) in Woods Hole from Tuesday November 29 to Friday December 2, 2016. The meeting was chaired by Dr Patrick J. Sullivan from the New England Fisheries Management Council's Scientific and Statistical Committee. The review panel (the Panel) was composed of three scientists appointed by the Center for Independent Experts: Vivian Haist, Neil Klaer, and Anders Nielsen. The SARC was assisted by the NEFSC Stock Assessment Workshop (SAW) Chairman, Dr James Weinberg, Ms Sheena Steiner, and Dr Chris Legault.

Draft stock assessment reports as well as all associated background documents were made available via a public web site to the Panel on 11 November for Witch flounder and 15 November for Black sea bass prior to the review meeting. During the meeting, all documents were available electronically via the same web site, and additional documents and presentations made during the meeting were also posted there.

1.2 Review Activities

A pre-review meeting on Tuesday morning of the Panel, Chair, Dr Weinberg and Dr Russell Brown (NEFSC) was held to discuss meeting logistics and reporting requirements. Time was taken after this meeting to assign draft reporting tasks among the Panel for working towards completion of a draft SARC 62 Summary Report by Friday. Terms of Reference (ToRs) were assigned according to broad categories: input data (N Klaer – ToRs 1-5 Black sea bass, ToRs 1-3 Witch flounder), stock assessment (V Haist – ToRs 5-8 Black sea bass, ToRs 4-6 Witch flounder), and projections and uncertainty (A Nielsen – ToRs 9-10 Black sea bass, ToRs 7-9 Witch flounder).

The meeting generally followed the draft agenda and included presentations by representatives of the stock assessment teams: Dr Gary Shepherd for Black sea bass and Dr Susan Wigley for Witch flounder. The meeting was open to the public and was available also on live audio via the web. Discussion was mostly restricted to clarification during presentations with broader discussion at the conclusion of presentations. Rapporteur notes were taken throughout. There was limited time available for additional work outside the meeting, but several homework analyses were requested by the Panel and results were provided during the meeting. The Panel participated in the review of each TOR for the meeting. Further drafting and agreement on the Assessment Summary Report for both species was completed on Thursday with particular assistance from Gary Shepherd, John Maniscalco, Susan Wigley, Mark Terceio, Liz Brooks and Chris Legault. On the final day, the meeting draft Summary Report was compiled from Panel drafts and edited, with additional editing in following weeks.

The meeting atmosphere was congenial, the venue was excellent, and contributors to the meeting all did so in a professional and efficient manner, allowing the meeting work to be completed on-time and without contention.

2 Review of assessments of Black sea bass and Witch flounder

The comments below refer to aspects that were examined during the meeting, but include my own additional commentary for preparation of this CIE report.

2.1 Findings by term of reference for Black sea bass

2.1.1 Summarize the conclusions of the February 2016 SSC peer review regarding the potential for spatial partitioning of the black sea bass stock. The consequences for the stock assessment will be addressed in TOR-6.)

This TOR was met satisfactorily. A Mid-Atlantic Fisheries Management Council Scientific and Statistical Committee peer review approved the SAW Working Group recommended spatial partitioning of the black sea bass population north of Cape Hatteras into North and South sub-units using the Hudson Canyon as the boundary. The Panel agrees that evidence, particularly from tagging and observed differences in recruitment patterns within sub-units, provides a good basis for this choice of sub-units. It was also noted that there is some degree of mixing among units particularly due to movement of some Northern fish along the shelf in winter and caught in the Southern unit.

2.1.2 Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch. Describe the spatial and temporal distribution of fishing effort.

This TOR was met satisfactorily. There were several components comprising both the commercial (bottom trawl, handline, pot and other) and the recreational fisheries and each was well described. Methods used for the calculation of discards and associated errors seem appropriate. Evidence for evaluation of errors in landings for commercial fleets should be assembled and quantified if possible. The major source of uncertainty influencing the stock assessment appears to derive from recreational catch, which is a significant component of the fishery, and the associated discard mortality. Fine-scale seasonal maps of trawl fishing catch by year and season provided a good way to demonstrate and visualize the spatial and temporal distribution of the commercial fishery catches.

Methods used to construct trawl and non-trawl catch at age matrices combined and by sub-unit appear appropriate given the available data. Further work is required to characterize uncertainty – either by assignment of error (at least annually), or through construction of alternative matrices that provide scenarios that account for that error for stock assessment sensitivity testing (see research recommendations).

2.1.3 Present the survey data being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of fishery dependent indices as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data. Evaluate model assumptions, estimates, and major sources of uncertainty.

This TOR was met satisfactorily. There are many abundance indices and associated data sources. To assist review, a summary table describing important aspects of each index should be provided in future (the requested table is shown here as Table 1). The Panel suggests that a review of the *a priori* degree of bias in abundance indices as a ranking be conducted by a set of technical experts that best understands the nature of the input data. Such information would provide guidance for how they are best included into a stock assessment (if at all). The recreational catch per angler index is fishery-dependent, covers a wide stock area and catch age distribution and is judged *a priori* as the index that best corresponds to the population level (this index was ranked 1st among considered indices – see Table 1).

Data filtering methods and index standardization procedures were not examined in detail for this review, but the information provided suggests that best-practice methods for these could be considered for review at a national level.

Data collection associated with the fishery-independent trawl indices provide a good source of length and age data, and lengths are available from all surveys.

As there are many indices that bridge the R/V Albatross/Bigelow change in the NEFSC index, a split of that index for this species seems justified. The NEFSC Albatross survey also covers a wide stock area and catch age distribution and was judged as the second most important index for modeling fitting purposes for this stock. There are very obvious differences in abundance trends in the North and South sub-units further justifying an area-based approach, as evidenced by the recreational catch per angler data (Figures A24 and A25 in the SAW62 Stock Assessment Report).

While there is long term agreement in trends between indices, there is some conflict in trends and peak years among indices for ages 1-8+ and age 1 within the North and South sub-units (e.g., VIMS, MD, DE age 1 indices in the S, Figures A9-A11 in the SAW62 Stock Assessment Report). To assist future reviews, plots of normalized comparable indices within each assessed sub-unit should be provided. There is general agreement that a recent increase occurs in abundance in the North sub-unit in both the NEFSC Bigelow and recreational CPA.

Table 1. Abundance index summary

Index Name	Assessed Area	Season	Location	Years	Fishery Independent	Type	Age Range	Filter Method	Derivation Method	Data	Spatial Range	Rank
NEFSC Winter Bottom Trawl Survey	S	Feb-Mar	Offshore	1992-2007	Yes	Absolute	1-8+	Strata subset	Stratified mean	Len,Wt & Age	Wide	3
NEFSC Spring Bottom Trawl Survey ALBATROSS	N + S	Mar-May	Offshore	1989-2008	Yes	Absolute	1-8+	Strata subset	Stratified mean	Len,Wt & Age	Wide	2
NEFSC Spring Bottom Trawl Survey BIGELOW	N + S	Mar-May	Offshore	2009-2015	Yes	Absolute	1-8+	Strata subset	Stratified mean	Len,Wt & Age	Wide	3
MA Resourse Assessment Trawl Survey	N	May	Inshore	1989-2015	Yes	Relative	1-8+	Strata subset	GLM	Length	Narrow	4
RI Seasonal Bottom Trawl Survey	N	Apr-May	Inshore	1989-2015	Yes	Relative	1-8+	All Strata	GLM	Length	Narrow	4
CT Long Island Sound Trawl Survey	N	May-Jun	Inshore	1989-2015	Yes	Relative	1-8+	All Strata	GLM	Length	Narrow	4
NY Small Mesh Trawl Survey	N	May-Jul	Inshore	1990-2015	Yes	Relative	1	Max length	GLM	Length	Narrow	5
NJ Ocean Trawl Survey	S	Jun	Inshore	1989-2015	Yes	Relative	1-8+	All Strata	GLM	Length	Medium	4
DE 16ft Trawl Survey	S	Apr-Jun	Inshore	1989-2015	Yes	Relative	1	Max length	Stratified mean	Length	Narrow	5
MD Coastal Bays 16ft Trawl Survey	S	May-Jun	Inshore	1989-2015	Yes	Relative	1	Max length	Stratified mean	Length	Narrow	5
VIMS Juvenile Fish and Blue Crab Trawl Survey	S	May-Jul	Inshore	1989-2015	Yes	Relative	1	Max length	Stratified mean	Length	Narrow	5
NEAMAP Ocean Trawl Survey	N + S	Apr-May	Coastal	2008-2015	Yes	Relative	1-8+ (N), 1 (S)	All Strata	Stratified mean	Len,Wt & Age	Medium	3
Recreational CPUE	N + S	Mar-Dec	Coastal	1989-2015	No	Relative	1-8+	Multi-Species effort	GLM	Length	Wide	1

Ranked by
Area coverage
Time series length
Ages represented

2.1.4 Consider the consequences of environmental factors on the estimates of abundance or relative indices derived from surveys.

This TOR was satisfactorily met. Results were presented from a study that concluded that warm saline conditions improved juvenile survival and the location of the shelf-slope front dictates the distribution of adults in winter offshore habitat. If this observation is to be used in assessment models, then additional research may be required to best support changes in trends. Further work is also encouraged to determine environmental factors affecting abundance of different life stages – particularly in an ecosystem context for the NW Atlantic region. Broad patterns of change across species complexes require characterization that may lead to co-variates useful to stock assessments.

Apparent northward movement of the center of population abundance (as potentially seen in the animation of NMFS spring BTS catch 1968-2014) is consistent with range movement associated with climate change as observed globally, so further work on range shift across species in the region and implications for stock assessment also seems warranted.

2.1.5 Investigate implications of hermaphroditic life history on stock assessment model. If possible, incorporate parameters to account for hermaphroditism.

This TOR was satisfactorily met. Results were presented of a simulation study of the effect of hermaphroditism appropriate to measures of spawning stock biomass (SSB) for stock assessment. The Panel agreed that use of combined male and female SSB was most appropriate for this species at this time.

2.1.6 Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock), using measures that are appropriate to the assessment model, for the time series (integrating results from ToRs-1,-4, & -5 as appropriate), and estimate their uncertainty. Include a historical retrospective analysis and past projection performance evaluation to allow a comparison with most recent assessment results.

This TOR was met satisfactorily and I agree with the comments in the summary report. The Panel accepted the SAW Working Group's proposed base case and concurred that it provided a credible basis for providing management advice.

ASAP is a forward-running age-structured assessment model from the NOAA toolbox that has a long history of use for stock assessments and is considered to be an appropriate method for species with compatible available data. The method has identical or similar input data requirements to VPA (catch-at-age, weight-at-age, abundance indices), has provided a good bridge from VPA to forward-running age-structured models, and has often been used for this purpose. Spatial sub-structuring is not explicitly handled, so if spatial sub-units are to be dealt with using ASAP, they are each modeled as separate independent populations.

As part of the data preparation process, complete catch-at-age matrices for each fleet (for landings and discards if required) are constructed. Various rules are used to complete regions of these matrices that have known catch but little or no associated age or length composition samples. Such rules as applied to Black sea bass appear appropriate.

There are good reasons (as also highlighted by previous reviews) to consider modeling the north and south sub-units separately – principally the apparent differences in abundance trends and appearance of strong year classes in composition data between the areas.

Application of the ASAP model to each sub-unit was largely successful (including an account of catch exchange between sub-units), except that strong and opposing retrospective patterns are evident within each sub-unit. To examine the entire stock, results from the assessed sub-units are added together, and when this is done, some retrospective pattern remains, although not sufficient to move the current stock status out of the not overfished and not overfishing quadrant after application of a rho adjustment.

Stock Synthesis (currently SS3) is also available from the NOAA toolbox and allows spatial sub-units within a single assessed population, and does not require the data preparation step of construction of comprehensive catch-at-age matrices. Although presented as a preliminary model for information only, an SS3 model was constructed that allowed partitioning of annual recruitment levels differently to each sub-area annually, and potentially allowed for mixing among areas. This model did not exhibit a substantial retrospective pattern (requested during this review) that would require rho adjustment. A comparison of overall SSB and recruitment trends of the SS3 model and the combined ASAP results for the north and south sub-units (requested during this review) showed overall agreement (Figure 1). This gave the Panel additional confidence in the combined ASAP model results.

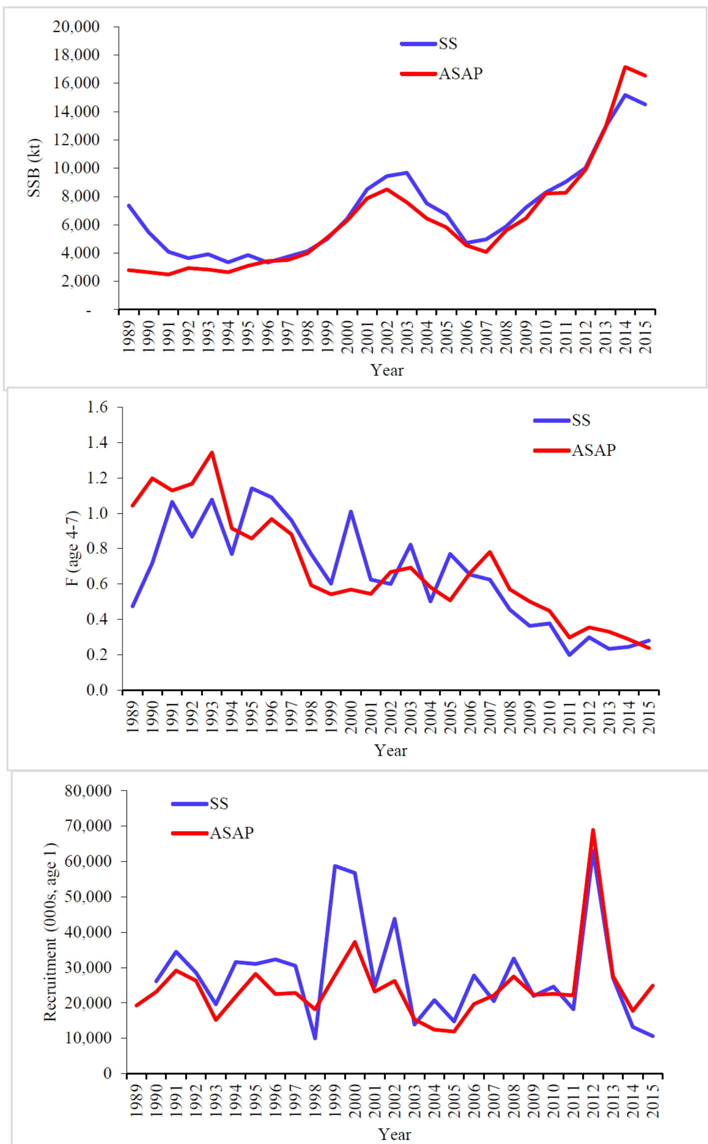

Figure 1 Comparison of Black sea bass ASAP two area model and SS (run 134)

Likelihood profiles for the value of M (assumed fixed at 0.4 for the base model) were given in Figure A164 of the Assessment Report. This shows a well-defined minimum at 0.4 for the combined area model, and near 0.2 for the north sub-unit and 0.8 for the south. A "Piner plot" that shows individual likelihood components for the combined area model was requested during the review (Figure 2) and shows (as normally expected) that different data sources contribute differently to the overall M likelihood profile. This shows that the relative weighting among the various input data sources would have an important influence on the likely value for M if estimated. The fixed value of M is still a major component of uncertainty for the Black sea bass assessment.

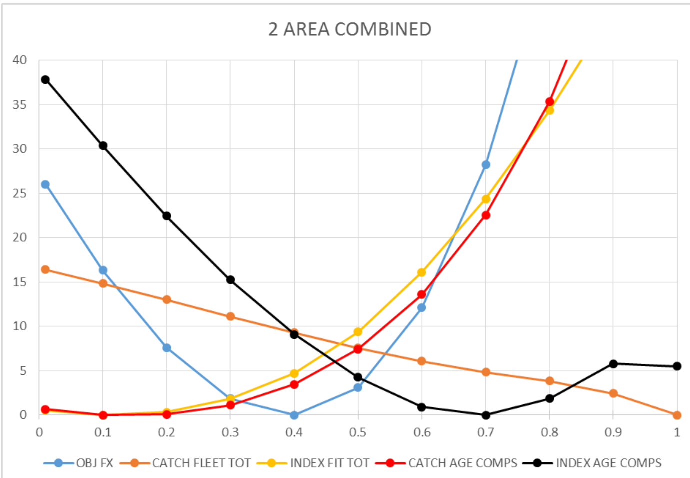

Figure 2 "Piner plot" of likelihood components of a profile on M for the two-area combined model

Weighting schemes as implemented seem supportable, but require further justification. Diagnostic plots of final weighted observations versus expected values for the indices do show that full iterative reweighting would lead to some changes in model results – e.g., north NEFSC index (Assessment Report Fig A81) expected values miss 95% confidence bounds for 5 points suggesting that the index CV needs to be increased, north REC CPA index (Assessment Report Fig A92) expected values fit within all confidence bounds suggesting that the index CVs might be reduced. Note however, that if CVs are directly from a statistical analysis of observation error they are not, by convention, normally reduced through reweighting. Recent work, particularly by Francis (e.g., Francis 2011) recommends relative down-weighting of composition versus abundance index data, and further work is required to investigate the implications of such advice for the Black sea bass assessment. Ideally, stock assessment packages would be self-weighting and not require such considerations, but such packages are not currently available, at least in the NOAA toolbox.

I agree with the approach of estimation of q values associated with surveys within the model and then making a comparison of the estimated and expected values as a test of model plausibility. This implies that there is an existing prior distribution for expected survey q values that is generally understood (often with an acceptable range of a factor of about 2). To further clarify the general understanding of survey q prior distributions, and to assist in making direct use of such priors in future assessments, further work is suggested to attempt to quantify and document all sources of uncertainty for survey q values and how they might contribute to an overall q prior distribution.

2.1.7 Estimate biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY}, and MSY), including defining BRPs for spatially explicit areas if appropriate, and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the appropriateness of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This TOR was met satisfactorily. Biological reference points were calculated using results of the two-area stock assessments. Reference points were calculated separately for the Northern and Southern sub-units, and biomass metrics summed and fishing mortality metrics averaged across the sub-units. Given uncertainty in the stock recruitment relationship, $F_{40\%}$ was chosen as a proxy for the F_{MSY} reference point and spawning stock biomass at $F_{40\%}$ ($SSB_{40\%}$) as the proxy for the stock biomass target reference point. Uncertainty in the BRPs, estimated using an MCMC algorithm, reflects uncertainty associated with the base assessment models but does not reflect the additional uncertainty associated with model misspecification or structural uncertainty. The Panel agrees the BRPs calculated for black sea bass are appropriate. No BRPs existed for this stock under the previous assessment.

Conveyance of model and projection uncertainty to management is currently built around the selection of a single best model, and the characterization of uncertainty with respect to that model, without accounting for model structural uncertainty. Model structural uncertainty is due to the necessarily simple assumptions in an assessment model relative to the more complex and unknowable complexity of the actual real-world fish population. Alternative simplifying assumptions made for the base case (e.g., fixing certain parameter values at point estimates, inclusion of particular abundance indices, selection of relative weighting for available data sets) can provide a range of plausible alternative values, and can result in a number of plausible alternative models. Provided that the relative plausibility of alternative models can be factored into weighting of the alternative models, a suite of plausibility-weighted models can provide information on the additional uncertainty not captured by the base case. Model structural uncertainty is normally large in comparison to within-model uncertainty as estimated by MCMC, etc., and should be considered when providing information about model uncertainty to management. Rice and Harley (2013) provide an example of a somewhat systematic exploration of model structural uncertainty and implications for stock status.

2.1.8 Evaluate overall stock status with respect to a new model or new models that considered spatial units developed for this peer review.

This TOR was met satisfactorily. The Panel supports the conclusion of the SAW that the black sea bass stock is not overfished and overfishing is not occurring. The Panel believed that the use of the rho-adjustment was reasonable because of the cancelling effect when area sub-units were combined and the adjustments did not influence the status of the stock determination when applied. The consistency among stock reconstructions from alternative formulations of the ASAP model and from an alternative model (SS3) suggest that these results will be robust to uncertainties associated to modeling assumptions.

2.1.9 Evaluate overall stock status with respect to a new model or new models that considered spatial units developed for this peer review.

a. **Provide numerical annual projections (3-5 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level) that fully incorporates observation, process and model uncertainty (see Appendix to the SAW ToRs). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and definition of BRPs for black sea bass).**

b. **Comment on which projections seem most realistic. Consider major uncertainties in the assessment as well as the sensitivity of the projections to various assumptions.**

c. **Describe this stock's vulnerability (see "Appendix to the SAW ToRs") to becoming overfished, and how this could affect the choice of ABC.**

The panel concluded that this TOR was completed satisfactorily. Short term projections were carried out for each area sub-unit and results were summed to get projections for the combined stock. This is consistent with the accepted model. The uncertainty in the final year's estimate is represented by an MCMC sample from its posterior distribution. The recruits used in the projections are sampled from the smoothed empirical distribution of estimated recruits in the period 2000-2015. The estimation uncertainty and recruitment process uncertainty are propagated through in the projections. The statistical distributions of the projected quantities are summarized by their means and standard deviations. We note that the TOR might be even better addressed if the probabilities of SSB below the threshold in the projected years were directly provided. Projections are carried out for two fishing scenarios, F at status quo and F at the proxy F_{MSY}. Model sensitivities were explored by comparing projections with and without rho-adjustment and a projection from the overall combined model. We note that the sensitivities could be further examined using short term projections supplied for the model sensitivity runs. The SAW report indicates that the rho-adjustment seems most realistic. The Panel notes that even though projections are conducted for each sub-unit, the combined projections should only be used, because of the major retrospective issues seen within each sub-unit.

2.1.10 Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

The panel concluded that this TOR was completed satisfactorily. The research recommendations from recent assessments are listed in the SAW report and progress on each recommendation is described. Further research recommendations from the SAW are put forward. The existing recommendations include: multiple age-structured models, species specific surveys, expand tagging, genetic studies, sex-change research, age reading validation, climate impacts, study catchability in gear types, investigate social and spawning dynamics, habitat studies, and evaluate use of industry samples. The SAW expressed concern about how to convince managers about their estimated uncertainty levels. To address this, the following two research recommendations are made: 1) Research into using self-weighting models. Uncertainty estimates of estimated quantities are obtained by propagating observation uncertainties through the models. When observation uncertainties are subjectively assigned, then so are the uncertainties of the results. 2) Application of prediction based methods to validate that actual prediction uncertainties correspond to estimated prediction uncertainties.

In addition, I include for consideration the following as potential additional areas of future research:

- Evidence for evaluation of errors in landings for commercial fleets should be assembled and quantified if possible. This may be accomplished either by assignment of annual error to catch, or through construction of alternative catch-at-age matrices that provide scenarios that account for that error for stock assessment sensitivity testing. These might be considered (for simplicity here) as components of model structural uncertainty.

- Data filtering methods and index standardization procedures were not examined in detail for this review, but the information provided suggests that best-practice methods for these could be considered for review at a national level. For example, there seems to be some inconsistency among various US stock assessment regions in filtering methods applied to similar or identical data sources (e.g., methods to select species “target” records from multi-species trips).

- Results were presented from a study that concluded that warm saline conditions improved juvenile survival and the location of the shelf-slope front dictates the distribution of adults in winter offshore habitat. If this observation is to be used in assessment models, then additional research may be required to best support changes in trends. Further work is also encouraged to determine environmental factors affecting abundance of different life stages – particularly in an ecosystem context for the NW Atlantic region. Broad patterns of change across species complexes require characterization that may lead to co-variates useful to stock assessments.

- Apparent northward movement of the center of population abundance is consistent with range movement associated with climate change as observed globally, so further work on range shift across species in the region and implications for stock assessment also seems warranted.

- Resolution of model bias as indicated by retrospective patterns requires the construction of alternative plausible hypotheses for the cause of such bias and the construction of corresponding alternative models. An approach that allows quick formulation of alternative models is to use a modeling framework that allows many additional flexibilities – particularly with regard to spatial sub-structuring or time-varying processes. An available framework with such flexibility is Stock Synthesis, and work to continue model exploration using such a flexible framework is encouraged.

- Recent work, particularly by Francis (e.g., Francis 2011) recommends relative down-weighting of composition versus abundance index data, and further work is required to investigate the implications of such advice for the Black sea bass assessment.

- To further clarify the general understanding of survey q prior distributions, and to assist in making direct use of such priors in future assessments, further work is suggested to attempt to quantify and document all sources of uncertainty for survey q values and how they might contribute to an overall q prior distribution.

- Conveyance of model and projection uncertainty to management is currently built around the selection of a single best model, and the characterization of uncertainty with respect to that model, without accounting for model structural uncertainty. Investigation of procedures for inclusion of structural uncertainty in management advice is recommended.

2.2 Findings by term of reference for Witch flounder

2.2.1 Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

The panel concluded that this TOR was generally completed satisfactorily as discussed below. The majority of the witch flounder catch was taken with otter trawl gear from the western Gulf of Maine and central basin and from deeper waters of the South Channel – principally in waters off the Maine and Massachusetts coast. A large decline in large market category fish occurred in the late 1980s. A reduction in catch over a long period from the Maine region appears to be heavily affected by changes in fisheries regulations (closures and catch restrictions). Methods for calculation of discards and associated errors seem appropriate. Evidence for evaluation of errors in landings for commercial fleet should be gathered and quantified if possible. Major sources of catch uncertainty are currently derived mainly from estimates of discards. However, while uncertainty in under-reported catch exists, it is not quantified. The magnitude of under-reported catch throughout the time series should be more fully assessed and documented. Using age-length keys on landings data from survey sources rather than those derived from landings age-length samples directly is problematic. If age samples are not available, it would be preferable to fit directly to the available length data by source.

2.2.2 Present available federal, state, and other survey data, indices of relative or absolute abundance, recruitment, etc. Characterize the uncertainty and any bias in these sources of data and compare survey coverage to locations of fishery catches. Select the surveys and indices for use in the assessment.

The Panel concluded that this TOR was completed satisfactorily. The spring and autumn NEFSC surveys are regarded as providing the best available fishery independent indices for this species and they show broadly similar patterns of a decline from the early 1960s to record low levels in the late 1980s and early 1990s, an increase to early 2000s followed by a declining trend. These surveys were combined across the Albatross-Bigelow transition period using a constant calibration factor. Pre-recruit indices are provided by the ASMFC and MENH surveys. Dealer Report LPUE, Vessel Trip Report, Observer Program and Study Fleet were evaluated as potential fishery-dependent indices. The Panel agrees with the evaluations made, and the provision of the LPUE index using 40% filtering as an appropriate best candidate. Many reasons for why such a series may be biased (e.g., under reporting) were given, although as noted during public comment, little quantified evidence to support the degree of bias was provided. The Panel agrees that the purpose of gathering fishery-independent series (such as the NEFSC) is to avoid such biases that are likely to be present in fishery-dependent indices. The Panel also agrees that quantification of bias in the LPUE index is a difficult technical problem and that quantitative measures of some important influencing factors may not be available. However, including LPUE in a sensitivity run, for example, might highlight how the fleet's perception of stock trends compares to model derived population trends, which might facilitate discussion with stakeholders. Perhaps including a time varying trend in selectivity and catchability for the fleet might allow this data to be usefully included into an assessment.

Data filtering methods and index standardization procedures were not examined in detail for this review, but information provided suggests that best-practice methods for these could be considered for review at a national level. This applies particularly to the potential inclusion of LPUE data, for example.

The sweep study provides a q estimate that can be applied to the NEFSC survey to determine absolute abundance and the Panel agrees that this is an appropriate piece of information to use for constructing an abundance series for this species. However, several assumptions are required to scale to the absolute values and care should be taken in evaluating these assumptions. For survey vessel calibration, consider including a prior representing the vessel comparison information when fitting these separate time series into the model or even including the calibration data explicitly into the model. Comparing the prior to the posterior estimates could prove fruitful in determining the efficacy of the calibration.

2.2.3 Investigate effects of environmental factors and climate change on recruitment, growth and natural mortality of witch flounder. If quantifiable relationships are identified, consider incorporating these into the stock assessment.

The panel concluded that this TOR was completed satisfactorily. Further work is encouraged to determine how environmental factors affect abundance at different life history stages – particularly in an ecosystem context for the NW Atlantic region. Broad patterns of change across species complexes require global characterizations that may lead to finding co-variates that are useful for informing stock assessments.

2.2.4 Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3 if appropriate), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections. Compare F's and SSB's that were projected during the previous assessment to their realized values.

This TOR was met. However, because the analytical assessment was rejected, reporting on some of these metrics is not possible. The Panel rejected the SAWs ASAP base model because the major retrospective pattern exhibited by the stock reconstruction was deemed unacceptable, and none of the sensitivity runs presented were considered to be an acceptable alternative as they also had major retrospective patterns or reflected unacceptable assumptions. Additionally, some of the alternative models had quite different abundance estimates or trends than the base model indicating results were not robust to the uncertain assumptions of the analysis.

Other problems with the assessment, that would in themselves not have been reason for rejecting the model, included: strong and common patterns in residuals (positive residuals for most of the final 15 years of the time series) for the fits to abundance survey indices

(NEFSC spring and autumn surveys and ASMFC survey) and catchability estimates of about 4 for the NEFSC surveys, which had an expectation of 1 based on the area swept study.

Retrospective analysis is a good method for identifying model misspecification (as caused by inconsistencies in data or model structure) and the apparent data quality issues or model misspecification in the witch flounder assessment needs to be resolved (e.g., applying a model or multiple models that remove the retrospective pattern while using an objective basis for model selection). Rho-based bias adjustment is an ad hoc procedure that may not correct such retrospective problems in the long term, and as such it may not always provide for appropriate management advice.

It is easy to say that further model structural exploration should be pursued in order to find alternatives that do not exhibit a strong retrospective pattern. I agree that this has already been fairly extensively attempted for Witch flounder – to an extreme of grid searching in multiple dimensions in terms of the extent of the retrospective bias. I have not seen work on retrospectives done more thoroughly elsewhere. However, several additional potential directions were shown by work provided as individual comment by Butterworth and Rademeyer – specifically the possibility of domed selectivity, the inclusion of the LPUE times series, and objectively-based down-weighting of the contribution to the likelihood of the CAA data (also examined separately as sensitivities for the ASAP model). There is often a concern that dome selectivity may lead to large levels of cryptic (usually spawning) biomass that can be difficult to justify without supporting evidence. An examination was requested during the review of the proportion of cryptic biomass generated by the Butterworth and Rademeyer reference case, and that proportion was significant at least at the start and end of the period assessed (Figure 3). Of course, a usual alternative to dome selectivity is to allow increased natural mortality at older ages, as also noted by Butterworth and Rademeyer. The main use of these alternative models for the review was however, to point out that it is possible to construct a number of alternative and potentially combined hypotheses and associated models that show at least less pronounced retrospective patterns and potentially improved fit to available data. Further work along these lines for Witch flounder is encouraged.

A known possible cause of retrospective patterns is the treatment of a time-varying process as time-invariant. I have participated in other reviews that have grappled with how far to go with the introduction of time-varying processes in assessment models, so I also add my own previous thoughts regarding that here for information:

> The only population biological parameter allowed to vary with time in most stock assessments is annual recruitment levels. Cumulative information on annual recruitment strength is provided fairly directly by composition data, so the reasons especially for high peaks and troughs in recruitment are usually apparent in the available data. It has also been recognized that other parameters are likely to vary through time – in particular natural mortality, but also growth and maturity. For natural mortality, it has been considered difficult to estimate time trends in changes without strong independent estimates for those changes, such as from ecosystem studies showing differences in predator abundance, and that time trends in M are difficult to disentangle from other factors such as catch mis-specification (e.g. see Brodziak et al., 2011). Allowing time variation in factors that directly affect productivity also leads to questions about choice of appropriate time periods for the selection of management reference points, and how to make appropriate stock projections.

Additional model parameters that may vary with time that are often dealt with using time-block methods are fishery/survey selectivity and catchability. For fisheries that are not associated with an abundance index, a fairly freely estimated time-varying pattern may be acceptable if it suitably captures annual fishery removals by size/age. For fishery-independent surveys the situation differs. Surveys are the most important source of abundance information for the model, particularly because at least the gear selectivity can be maintained as a constant through time. Availability (either by age or year) is another matter, but is usually treated as a source of additional random error. If a true trend (or even a step) exists in either survey selectivity or catchability, then that survey is biased, and the bias needs to be accounted for, or the survey truncated, split or discarded. Such a bias would ideally be investigated and identified with a focused study and auxiliary data not necessarily used in the assessment model. Adding annual time-variability to survey selectivity or catchability and finding that trends are estimated may simply be providing a means for the model to trade trends in population abundance to improve the fit to noisy composition data in preference to abundance indices. The reasons that such a model might result in trends in survey selectivity or catchability are not readily apparent from standard input data sources, and may be difficult to diagnose.

My own recommendation for now is that time variability should be allowed in a parameter when there is an available reliable data source that fairly directly measures such a change, and that a trend exists in that data source that needs to be captured by the assessment model.

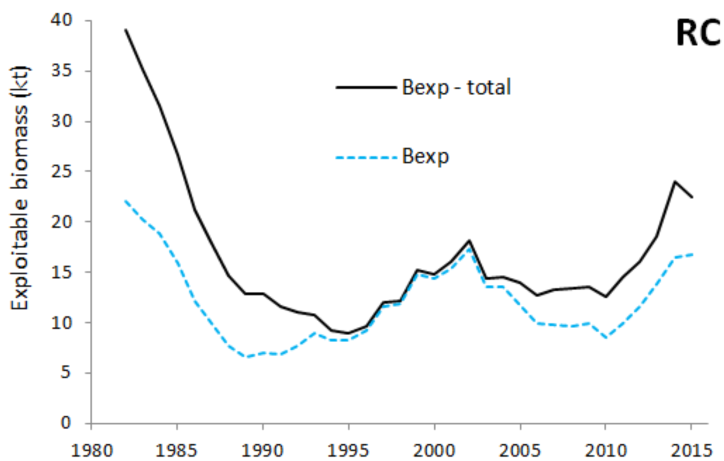

Figure 3 Trajectories of exploitable biomass (based on the 2005-2015 commercial selectivity) with and without the "cryptic" component for the reference case.

The panel believes that the previously accepted VPA model is not an acceptable alternative to the rejected ASAP application because it exhibits a similar major retrospective pattern. The ASAP sensitivity analyses evaluated one-off changes that included domed-shaped selectivity in either the fishery and survey, the fishery-dependent LPUE time series, and down-weighting the age-composition data series fitted in the model (Francis weights). The assumption of domed selectivity had minimal effect on model runs, likely because these were conducted separately for the fishery and survey (NEFSC) indices. While inclusion of the LPUE series and down-weighting age-composition data resulted in somewhat higher abundance estimates, these runs did not resolve the retrospective pattern, nor did they change the problematic survey residual patterns, or provide q estimates close to their expectation of 1.

Additional runs that examined alternative values for M or catch multipliers indicated that a 2-3 fold increases in M or a very large increase in catch were required to remove the retrospective pattern and were considered implausible. An SCAA model analysis was considered during the SAW and further analyses were provided during the public comment period of the SARC peer review meeting. This work helped identify other model configurations that might be usefully explored in the future provided that the assumptions used in creating them are valid.

The Panel recommends that the NEFMC SSC consider using the empirical approach discussed in the assessment document for use as the basis for developing management advice. The empirical analysis indicates stock biomass declined since 2002 although appears to have stabilized in recent years.

Because both the VPA and ASAP age-structured model applications were rejected, there is no basis for conducting a historical retrospective analysis, or to compare F's and SSB's that were estimated or projected during previous assessments.

2.2.5 State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model- based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

This TOR was met. However, because the analytical assessment was rejected, reporting on some of these metrics is not possible. Because of major retrospective patterns, the Panel rejected the analytical assessment conducted for the SARC review as well as the previous VPA model, so there is no basis for calculating model-based reference points.

Regarding the biological reference points estimated in the 2015 assessment update, these reference points were based on results of a VPA model that was rejected during the SARC 62 peer review process. The Panel supports the SAWs proposed alternative to use an F_{MSY} proxy of a relative exploitation of 0.05 in the near term, where other alternatives are unavailable. The basis of this value is ad hoc, calculated as the average of the most recent 9 years of estimated relative exploitation rates from the empirical approach. Over that period stock abundance was relatively stable. The empirical approach, based on the NEFSCs spring and autumn surveys, calculates swept-area abundance incorporating results of the swept-area experiment. While absolute biomass estimates from this method will be highly uncertain (because of assumptions required in the conversion of survey catch rate to absolute estimates), the estimates of relative exploitation rates will be more robust to those uncertainties. The SARC did not have time to fully review this approach in comparison to other data poor procedures or to the Terms of Reference provided for the meeting.

The Panel does not recommend using the $F_{40\%}$ approach for catch advice because the assessment basis for that value has been rejected.

2.2.6 Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model (or possibly models, in accord with guidance in attached "Appendix to the SAW Assessment ToRs") developed for this peer review. In both cases, evaluate whether the stock is rebuilt.

a. **When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the updated BRP estimates.**

   The Panel rejected the existing VPA model with current data for the same reasons discussed for the other age-structured analyses discussed above, so there is no basis for evaluating stock status relative to updated BRP estimates.

b. **Then use the newly proposed model (or possibly models, in accord with guidance in "Appendix to the SAW Assessment ToRs") and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-5).**

   The assessment model application was rejected due to a major retrospective pattern; therefore, an alternative empirical approach was used to determined biomass and exploitation rates, but not biological reference points. As a consequence, stock status is currently unknown.

2.2.7 Develop approaches and apply them to conduct stock projections.

a. **Provide numerical annual projections (3 years) and the statistical distribution (e.g., probability density function) of the catch at FMSY or an FMSY proxy (i.e. the overfishing level, OFL) (see Appendix). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, magnitude and variability in recruitment).**

b. **Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions. Identify reasonable projection parameters (recruitment, weight-at- age, retrospective adjustments, etc.) to use when setting specifications.**

c. **Describe this stock's vulnerability to becoming overfished, and how this could affect the choice of ABC. The choice takes scientific uncertainty into account (see Appendix).**

This TOR was met. However, because the analytical assessment was rejected reporting on some of these metrics is not possible. The alternative empirical area-swept method does not offer projections, as it contains no process model for the stock dynamic, which is needed for short-term and long-term projections. The empirical area-swept method does not provide a biomass threshold, but does indicate that the stock is at low historical levels.

2.3.9 Review, evaluate and report on the status of research recommendations from the last peer reviewed benchmark stock assessment. Identify new research recommendations.

The panel concluded that this TOR was completed satisfactorily. The research recommendations from previous benchmarks are listed in the SAW assessment report and progress on each recommendation is described. Further research recommendations from the SAW are put forward. Existing recommendations include: refining calibration factors, examine mean weight trends, research in causes for retrospective patterns, aging archived samples, stock identification, tagging, larval index, environmental/habitat preferences, influence of age-composition data, spatial modelling, and investigate plausible M changes.

The panel noted and support that many of the recommendations relate to solving the major retrospective issue. This issue is seen in many stocks in the region, so any insight gained here would be widely beneficial. The panel recommends that this issue be addressed as a research track topic. Focus should be on identifying causes that could lead to such retrospective patterns, and then on evaluating how plausible each potential cause is. A list of scenarios to consider could include: time evolving or mis-specified: catchability, selectivity, natural mortality, misreporting, or age assignment. The scenarios could be constructed via simulations to validate that they could cause such retrospective patterns. When evaluating how plausible each scenario is for the real data, it may be useful to run prediction-based validations (estimate from one part of the data and predict the remaining). The panel is aware that this is a large undertaking and assessment history in the region shows that no quick fixes should be expected.

In addition, I include for consideration the following as potential additional areas of future research:

- Evidence for evaluation of errors in landings for commercial fleets should be assembled and quantified if possible. This may be accomplished either by assignment of annual error to catch, or through construction of alternative catch-at-age matrices that provide scenarios that account for that error for stock assessment sensitivity testing. These might be considered (for simplicity here) as components of model structural uncertainty.

- Data filtering methods and index standardization procedures were not examined in detail for this review, but the information provided suggests that best-practice methods for these could be considered for review at a national level. For example, there seems to be some inconsistency among various US stock assessment regions in filtering methods applied to similar or identical data sources (e.g., methods to select species “target” records from multi-species trips).

- Further work is encouraged to determine environmental factors affecting abundance of different life stages – particularly in an ecosystem context for the NW Atlantic region. Broad patterns of change across species complexes require characterization that may lead to co-variates useful to stock assessments.

- Resolution of model bias as indicated by retrospective patterns requires the construction of alternative plausible hypotheses for the cause of such bias and the construction of corresponding alternative models. An approach that allows quick formulation of alternative models is to use a modeling framework that allows many additional flexibilities – particularly with regard to spatial sub-structuring or time-varying processes. An available framework with such flexibility is Stock Synthesis, and work to continue model exploration using such a flexible framework is encouraged.

- To further clarify the general understanding of survey q prior distributions, and to assist in making direct use of such priors in future assessments, further work is suggested to attempt to quantify and document all sources of uncertainty for survey q values and how they might contribute to an overall q prior distribution.

- Conveyance of model and projection uncertainty to management is currently built around the selection of a single best model, and the characterization of uncertainty with respect to that model, without accounting for model structural uncertainty. Investigation of procedures for inclusion of structural uncertainty in management advice is recommended.

References

Brodziak, J., Ianelli, J., Lorenzen, K. and Methot, R.D. 2011. Estimating Natural Mortality in Stock Assessment Applications. NOAA Technical Memorandum NMFS-F/SPO-199.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1124–1138.

Rice J. and S. Harley. 2013. Updated stock assessment of silky sharks in the western and central Pacific Ocean. Paper WCPFC-SC9-2013/ SA-WP-03 to the Ninth Regular Session of the Scientific Committee, Western and Central Pacific Fisheries Committee, Pohnpei, Federated States of Micronesia, 6-14 August 2013. 71 pp.

Annex 1: Bibliography of materials provided for review

Black Sea Bass

Background Papers

Blaylock J, Shepherd GR. 2016. Evaluating the vulnerability of an atypical protogynous hermaphrodite to fishery exploitation: results from a population model for black sea bass (*Centropristis striata*). Fish Bull. 114:476–489.

Brooks EN et al. 2008. Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points. Fish Bull. 106:12–23.

Keigwin B, Shepherd GR, Wuenschel MJ. 2016. Geomorphometric analysis indicates overlap in body shape between sexes of black sea bass (*Centropristis striata*). US Dept Commer, Northeast Fish Sci Cent Ref Doc. 16-07; 26p.

Miller AS, Shepherd GR, Fratantoni PS. 2016. Offshore Habitat Preference of Overwintering Juvenile and Adult Black Sea Bass, *Centropristis striata*, and the Relationship to Year-Class Success. PLoS ONE 11(1): 19p.

Moser J, Shepherd GR. 2009. Seasonal Distribution and Movement of Black Sea Bass (*Centropristis striata*) in the Northwest Atlantic as Determined from a Mark-Recapture Experiment. J Northw Atl Fish Sci. 40: 17–28.

Nieland JL, Shepherd GR. 2011. Comparing Black Sea Bass Catch and Presence Between Smooth and Structured Habitat in Northeast Fisheries Science Center Spring Bottom Trawl Surveys (Working Paper for SAW 53). 7p.

Northeast Fisheries Science Center. 2012. 53rd Northeast Regional Stock Assessment Workshop (53rd SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-03; 33p.

Shepherd, G., K. Shertzer, J. Coakley, and M. Caldwell (Editors). 2013. Proceedings from a workshop on modeling protogynous hermaphrodite fishes. Raleigh, NC. 33p.

Working Papers

Fay G. 2016. Retrospective analysis for Black sea bass Stock Synthesis model 'run_164'. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 6p.

Fay, G and Cadrin S. 2016. Simulation testing assessment models for Black Sea Bass. Appendix A6, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 26p.

Fay G, McNamee J, Cadrin S. 2016. Stock Synthesis Application to Black Sea Bass. Appendix A9, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 59p.

Fay G, McNamee J, Cadrin S. 2016. Stock Synthesis Application to Black Sea Bass. Appendix A9, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 59p.

Robillard E et al. 2016. Validation of Black Sea bass, Centropristis striata, Ages Using Oxytetracycline Marking and Scale Margin Increments. Appendix A1, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 18p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 247p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Summary Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 8p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Port-Based Black Sea Bass Outreach Project. Appendix A2, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 21p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Investigating the utility of inshore trawl surveys for developing black sea bass abundance indices. Appendix A3, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Fishery Management History. Appendix A4, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Term of Reference 1 – Spatial Issues. Appendix A5, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 46p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black sea bass distribution maps – Distribution of State and Federal surveys and NEFSC spring survey distribution maps, 1989-2015. Appendix A7, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 206p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. ALK simulation to test efficacy of multinomial approach. Appendix A8, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 19p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Overall Model of All Plots. Appendix A10, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 156p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. North Model of All Plots. Appendix A11, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 134p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. South Model of All Plots. Appendix A12, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 116p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. North Area Exchange of All Plots. Appendix A13, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 134p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. South Area Exchange of All Plots. Appendix A14, Stock Assessment Report of Black Sea Bass. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 116p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Two Area Model Justification. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Combined ASAP Retros. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 3 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Comparison of results for Black Sea Bass ASAP Two Area model and SS (run 134). SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Groundfish retro-adjusted values used in management. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Index tables. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. M Profile Obj FX Components. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Model Justification Diagnostics. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 8 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Normalized indices used in both North and South area models. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 4 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. SS comparisons. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 1p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Standardized Age Comp Residual Plots. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 18 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock recruit. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 2 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black sea bass Z-score normalized index values. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 1 slide.

Presentations

Working Group, Stock Assessment Workshop (SAW 62). 2016. Black Sea Bass Assessment Review. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 261 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Commercial VTRs. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 40 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. VTR Trawl and Spring Survey. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 23 slides.

Witch Flounder

Background Papers

Butterworth DS and Rademeyer RA. 2016. Further Remarks on Gulf of Maine-Georges Bank Witch Flounder Assessment Results. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 12p.

Cadrin S and Wright B. 2016. Fishery Catch Rates of Working Flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 17p.

DeCelles G. 2016. An Assessment of Witch Flounder (*Glyptocephalus cynoglossus*) Stock Structure. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 18p.

Friedland K. 2016. Data to inform habitat model construction for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 23p.

Friedland K. 2016. Estimated witch flounder habitat using random forest models. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Hare J et al. 2016. In situ temperature and salinity data for use in stock assessments. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Hare J et al. 2016. Empirical estimates of maximum catchability of Witch Flounder *Glyptocaphalus cynoglossus L.* on the Northeast Fisheries Science Center Fall bottom trawl survey. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Hare J et al. 2016. Environmentally explicit stock-recruitment relationships in Witch Flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 5p.

Kritzer JP et al. 2016. Spatial and Temporal Patterns in Habitat Use and Depth Distribution of Witch Flounder: Implications for Stock Assessment. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 11p.

Northeast Fisheries Science Center. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dep Commer, NOAA Fisheries, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii.

Northeast Fisheries Science Center. 2012. Assessment or Data Updates of 13 Northeast Groundfish Stocks through 2010. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-06; 789 p.

Northeast Fisheries Science Center. 2015. Operational Assessment of 20 Northeast Groundfish Stocks, Updated Through 2014. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 15-24; 251 p.

Odell J et al. 2016. NSC-AFM-GFCPF Witch Flounder Letter. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 2p.

Palmer MC. 2016. Catch curve analysis of witch flounder fishery and survey catch-at-age data. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 9p.

Richardson D. 2016. A minimum estimate of Witch Flounder spawning stock biomass using experimental estimates of catchability on the NEFSC trawl survey. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Terceiro M. 2016. TOR 1: Description of commercial fishery Dealer Report trawl gear landings and effort and modeling landings rate (LPUE) data for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 21p.

Terceiro M. 2016. TOR 1: Description of commercial fishery Dealer Report trawl gear landings and effort and modeling landings rate (LPUE) data for witch flounder: 'Directed' Trips (=>40% of trip landings). Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 11p.

Terceiro M. 2016. TOR 1 & 2: Modeling commercial fishery Dealer Report fish trawl gear landings rate (LPUE) data for witch flounder: 'Directed' Trips (=>40%, =>25%, and =>10% of trip landings). Working

Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 27p.

Terceiro M. 2016. TOR 1: Description and modeling of NEFOP (Observer) fish trawl gear catch rate (CPUE) data for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 17p.

Terceiro M. 2016. TOR 1 & 2: Description of Vessel Trip Report trawl gear catch and effort data and modeling catch rates (CPUE) for witch flounder. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 30p.

Walsh HJ et al. 2016. Changes in the distributions of larval, juvenile, and adult witch flounder in the Northeast US Shelf Ecosystem: Updates Through 2015. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 9p.

Wigley SE. 2016. Rough vs Smooth Bottom Type: An Initial Exploration. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 15p.

Wigley SE and Burnett JM. 2016. Preliminary Estimates of Biological and Yield Characteristics of Deep-water Witch Flounder (Glyptocephalus cynoglossus) in the Georges Bank-Mid-Atlantic Bight Region. J Northw Atl Fish Sci. 31:181-194.

Wigley SE. 2016. Refinements to 1982-2014 Witch Flounder Discard Estimates. Working Paper for SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 39p.

Wigley SE, Brodziak JKT,Col L. 2003. Assessment of the Gulf of Maine and Georges Bank witch flounder stock for 2003. Northeast Fish. Sci. Cent. Ref. Doc. 03-14; 186 p.

Working Papers

Butterworth DS and Rademeyer RA. 2016. Response to reviewer requests in regard to the impact of selectivity doming in the preferred SCAA model. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 4p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment of Witch Flounder for 2016. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 523p.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Stock Assessment Summary of Witch Flounder for 2016. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 16p.

Presentations

Hare J et al. 2016. Empirical Estimates of Maximum Catchability of Witch Flounder on the Northeast Fisheries Science Center Fall Bottom Trawl Survey. SAW/SARC 62. November 29 – December 2, 2016.

NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 23 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder Assessment Review, TORs 1-3. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 89 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder Assessment Review, TORs 4-9. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 123 slides.

Working Group, Stock Assessment Workshop (SAW 62). 2016. Witch Flounder SARC Discussion Slides. SAW/SARC 62. November 29 – December 2, 2016. NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. Power Point presentation. 32 slides.

Annex 2: Copy of the Statement of Work

Statement of Work
National Oceanic and Atmospheric Administration (NOAA) National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program External Independent Peer Review

62nd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC) Benchmark stock assessment for Black sea bass and Witch flounder

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards.
(http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information may be obtained from www.ciereviews.org.

Scope

The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC peer review is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development and report preparation (which is done by SAW Working Groups or ASMFC technical committees), assessment peer review (by the SARC), public presentations, and document publication. This review determines whether or not the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fisheries within the jurisdiction of NOAA's Greater Atlantic Regional Fisheries Office (GARFO).

The purpose of this meeting will be to provide an external peer review of a benchmark stock assessment for **Black sea bass and Witch flounder**. The requirements for the peer review follow. This Statement of Work (SOW) also includes Appendix 1: TORs for the stock assessment, which are the responsibility of the analysts; Appendix 2: a draft meeting agenda; Appendix 3: Individual Independent Review Report Requirements; and Appendix 4: SARC Summary Report Requirements.

Requirements
NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The SARC chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the SARC chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the SOW, OMB Guidelines, and the TORs below. All TORs must be addressed in each reviewer's report. No more than one of the reviewers selected for this review is permitted to have served on a SARC panel that reviewed this same species in the past. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include forward projecting statistical catch-at-age models. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points (BRPs) that includes an appreciation for the varying quality and quantity of data available to support estimation of BRPs. For Black sea bass, knowledge of spatial models and complex fisheries with multiple fleets and recreational fisheries would be useful. For Witch flounder, knowledge of flatfish ecology would be useful.

Requirements for Reviewers

- Review the background materials and reports prior to the review meeting
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus.
- Each reviewer shall assist the SARC Chair with contributions to the SARC Summary Report
- Deliver individual Independent Review Reports to the Government according to the specified milestone dates

- This report should explain whether each stock assessment Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified below in the "Requirements for SARC panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.
- The Independent Report can also be used to provide greater detail than the SARC Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Requirements for SARC panel

- During the SARC meeting, the panel is to determine whether each stock assessment Term of Reference (TOR) of the SAW was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the SARC chair shall identify or facilitate agreement among the reviewers for each stock assessment TOR of the SAW.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the SOW and Schedule of Milestones and Deliverables below.

Requirements for SARC chair and reviewers combined:

Review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

The SARC Chair, with the assistance from the reviewers, will write the SARC Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each stock assessment Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair's objective during this SARC Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair's opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion. The SARC Summary Report will not be submitted, reviewed, or approved by the Contractor.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

Foreign National Security Clearance
When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, country of birth, country of citizenship, country of permanent residence, country of current residence, dual citizenship (yes, no), passport number, country of passport, travel dates.) to the NEFSC SAW Chair for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and at the Northeast Fisheries Science Center in Woods Hole, Massachusetts.

Period of Performance
The period of performance shall be from the time of award through January 15, 2017. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

No later than November 15, 2016	Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
No later than November 15, 2016	NMFS Project Contact will provide reviewers the pre-review documents
Nov. 29 – Dec. 2, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA
December 2, 2016	SARC Chair and reviewers work at drafting reports during meeting at Woods Hole, MA, USA
December 16, 2016	Reviewers submit draft independent peer review reports to the contractor's technical team for review
December 16, 2016	Draft of SARC Summary Report, reviewed by all reviewers, due to the SARC Chair *
December 23, 2016	SARC Chair sends Final SARC Summary Report, approved by reviewers, to NMFS Project contact (i.e., SAW Chairman)
December 30, 2016	Contractor submits independent peer review reports to the COR and technical point of contact (POC)
January 6, 2017	The COR and/or technical POC distributes the final reports to the NMFS Project Contact and regional Center Director

* The SARC Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content; (2) The reports shall address each TOR as specified; (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $20,000.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

Project Contacts
Dr. James Weinberg, NEFSC SAW Chair Northeast
Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov Phone: 508-495-2352

Dr. William Karp, NEFSC Science Director
Northeast Fisheries Science Center 166 Water St.,
Woods Hole, MA 02543
william.karp@noaa.gov Phone: 508-495-2233

Appendix 1. Terms of Reference for the SAW Working Group (62nd SAW/SARC Stock Assessment)

T*he SARC Review Panel shall assess whether or not the SAW Working Group has reasonably and satisfactorily completed the following actions.*

A. Black sea bass

1. Summarize the conclusions of the February 2016 SSC peer review regarding the potential for spatial partitioning of the black sea bass stock. The consequences for the stock assessment will be addressed in TOR-6.)

2. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data. Evaluate available information on discard mortality and, if appropriate, update mortality rates applied to discard components of the catch. Describe the spatial and temporal distribution of fishing effort.

3. Present the survey data being used in the assessment (e.g., indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of fishery dependent indices as a measure of relative abundance. Characterize the uncertainty and any bias in these sources of data.

4. Consider the consequences of environmental factors on the estimates of abundance or relative indices derived from surveys.

5. Investigate implications of hermaphroditic life history on stock assessment model. If possible, incorporate parameters to account for hermaphroditism.

6. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock), using measures that are appropriate to the assessment model, for the time series (integrating results from TORs-1,-4, & -5 as appropriate), and estimate their uncertainty. Include a historical retrospective analysis and past projection performance evaluation to allow a comparison with most recent assessment results.

7. Estimate biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY}, and MSY), including defining BRPs for spatially explicit areas if appropriate, and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the appropriateness of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

8. Evaluate overall stock status with respect to a new model or new models that considered spatial units developed for this peer review.

9. Develop approaches and apply them to conduct stock projections.
    a. Provide numerical annual projections (3-5 years) and the statistical distribution (e.g., probability density function) of the OFL (overfishing level) that fully incorporates observation, process and model uncertainty (see Appendix to the SAW TORs). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and definition of BRPs for black sea bass).
    b. Comment on which projections seem most realistic. Consider major uncertainties in the assessment as well as the sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

10. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

B. Witch flounder

1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

2. Present available federal, state, and other survey data, indices of relative or absolute abundance, recruitment, etc. Characterize the uncertainty and any bias in these sources of data and compare survey coverage to locations of fishery catches. Select the surveys and indices for use in the assessment.

3. Investigate effects of environmental factors and climate change on recruitment, growth and natural mortality of witch flounder. If quantifiable relationships are identified, consider incorporating these into the stock assessment.

4. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-3 if appropriate), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections. Compare F's and SSB's that were projected during the previous assessment to their realized values.

5. State the existing stock status definitions for “overfished” and “overfishing”. Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the “new” (i.e., updated, redefined, or alternative) BRPs.

6. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model (or possibly models, in accord with guidance in attached “Appendix to the SAW Assessment TORs”) developed for this peer review. In both cases, evaluate whether the stock is rebuilt .
    a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the updated BRP estimates.
    b. Then use the newly proposed model (or possibly models, in accord with guidance in “Appendix to the SAW Assessment TORs”) and evaluate stock status with respect to “new” BRPs and their estimates (from TOR-5).

7. Develop approaches and apply them to conduct stock projections.
    a. Provide numerical annual projections (3 years) and the statistical distribution (e.g., probability density function) of the catch at F_{MSY} or an F_{MSY} proxy (i.e. the overfishing level, OFL) (see Appendix). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, magnitude and variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions. Identify reasonable projection parameters (recruitment, weight-at-age, retrospective adjustments, etc.) to use when setting specifications.
    c. Describe this stock’s vulnerability to becoming overfished, and how this could affect the choice of ABC. The choice takes scientific uncertainty into account (see Appendix).

8. Evaluate the validity of the current stock definition, taking into account what is known about migration, and make a recommendation about whether there is a need to modify the current stock definition for future stock assessments.

9. Review, evaluate and report on the status of research recommendations from the last peer reviewed benchmark stock assessment. Identify new research recommendations.

Clarification of Terms
used in the SAW/SARC Terms of Reference

Guidance to SAW WG about "Number of Models to include in the Assessment Report": In general, for any TOR in which one or more models are explored by the WG, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the WG and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Participation among members of a Stock Assessment Working Group:

Anyone participating in SAW meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Appendix 2. Draft Review Meeting Agenda

{Final Meeting agenda to be provided at time of award}

62nd Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC)
Benchmark stock assessment for A. Black sea bass and B. Witch flounder

Nov. 29 – Dec. 2, 2016

Stephen H. Clark Conference Room – Northeast Fisheries Science Center Woods Hole, Massachusetts

DRAFT AGENDA*

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Tuesday, Nov. 29

10 – 10:30 AM
Welcome — **James Weinberg**, SAW Chair
Introduction — **Pat Sullivan,** SARC Chair
Agenda
Conduct of Meeting

10:30 – 12:30 PM Assessment Presentation (A. Black sea bass)
Gary Shepherd **TBD**

12:30 – 1:30 PM Lunch

1:30 – 3:30 PM Assessment Presentation (A. Black sea bass)
Gary Shepherd **TBD**

3:30 – 3:45 PM Break

3:45 – 5:45 PM SARC Discussion w/ Presenters (A. Black sea bass)
Pat Sullivan, SARC Chair **TBD**

5:45 – 6 PM Public Comments

TOPIC PRESENTER(S) SARC LEADER RAPPORTEUR

Wednesday, Nov. 30

8:30 – 10:30 AM Assessment Presentation (B. Witch flounder)
Susan Wigley **TBD**

10:30 – 10:45 AM Break

10:45 – 12:30 PM Assessment Presentation (B. Witch flounder)
Susan Wigley **TBD**

12:30 – 1:30 PM Lunch

1:30 – 3:30 PM SARC Discussion w/presenters (B. Witch flounder)
Pat Sullivan, SARC Chair **TBD**

3:30 – 3:45 PM Public Comments

3:45 -4 PM Break

4 – 6 PM Revisit with Presenters (A. Black sea bass)
Pat Sullivan, SARC Chair **TBD**

7 PM **(**Social Gathering**)**

TOPIC	PRESENTER(S)	SARC LEADER	RAPPORTEUR

Thursday, Dec. 1

8:30 – 10:30 Revisit with Presenters (B. Witch flounder)
Pat Sullivan, SARC Chair **TBD**

10:30 – 10:45 Break

10:45 – 12:15 Review/Edit Assessment Summary Report (A. Black sea bass)
Pat Sullivan, SARC Chair **TBD**

12:15 – 1:15 PM Lunch

1:15 – 2:45 PM (cont.) Edit Assessment Summary Report (A. Black sea bass)
Pat Sullivan, SARC Chair **TBD**

2:45 – 3 PM Break

3 – 6 PM Review/edit Assessment Summary Report (B. Witch flounder)
Pat Sullivan, SARC Chair **TBD**

Friday, Dec. 2

9:00 AM – 5:00 PM SARC Report writing **

*All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public.
**During the SARC Report writing stage, the public should not engage in discussion with the SARC.

Appendix 3. Individual Independent Peer Review Report Requirements

1. The independent peer review report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the SARC Summary Report.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the SARC Summary Report that they believe might require further clarification.

   d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Statement of Work
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Appendix 4. SARC Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether or not each Term of Reference of the SAW Working Group was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

To make this determination, the SARC chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the SAW, and relevant papers cited in the SARC Summary Report, along with a copy of the Statement of Work.

The report shall also include as a separate appendix the assessment Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Annex 3: List of participants

SAW-SARC 62 ATTENDEES
NOVEMBER 29 – DECEMBER 2, 2016

NAME	AFFILIATION
James Weinberg	NEFSC
Russ Brown	NEFSC
Gary Shepherd	NEFSC
Mark Terceiro	NEFSC
Susan Wigley	NEFSC
Tony Wood	NEFSC
Kiersten Curti	NEFSC
Sheena Steiner	NEFSC
Chris Legault	NEFSC
Sarah Gaichas	NEFSC
Alicia Miller	NEFSC
Paul Nitzschke	NEFSC
Chuck Adams	NEFSC
Loretta O'Brien	NEFSC
John Maniscalco	NYDEC
Jamie Cournane	NEFMC
Kiley Dancy	MAFMC
Patricia Perez	NEFOP
Steve Cadrin	SMAST
Vivian Haist	Centre for Independent Experts (Reviewer)
Anders Nielsen	Centre for Independent Experts (Reviewer)
Neil Klaer	Centre for Independent Experts (Reviewer)
Pat Sullivan	NEFMC (SARC Chair)
Kirby Rootes-Murdy	ASMFC
Rich McBride	NEFSC
Nichola Meserve	MA DMF
Ashley Weston	SMAST
Greg Power	GARFO
Liz Daskey	Cornell Univ
Bob Glenn	MA DMF
Larry Alade	NEFSC
Mike Radziszewski	NEFOP
Brandon Muffley	MAFMC
Katherine Sosebee	NEFSC
Heath Cook	NEFSC
Tim Miller	NEFSC
Gavin Fay	UMASS Dartmouth
Greg DeCelles	MA DMF
Jackie Odell	NSC

SAW-SARC 62 ATTENDEES – Page 2

NAME	AFFILIATION
Vito Giacalone	NSC
Liz Brooks	NEFSC
Brooke Wright	SMAST
Aja Szumylo	GARFO
Dave Richardson	NEFSC
Harvey Walsh	NEFSC
Amanda Hart	UMASS Dartmouth
Hannah Goodale	GARFO
Jessica Blaylock	NEFSC
Melanie Griffin	